

## **Identifying refugees and other migrant groups in European large-scale surveys: An explorative analysis of integration outcomes by age upon arrival, reasons for migration and country of birth groups using the European Union Labour Force Survey 2014 Ad Hoc Module**

### **Abstract**

**Aims:** To explore the association between self-reported reasons for migration, age upon arrival and Eurostat's country of birth classification, and to study these measures in relation to education, employment and language-skills.

**Methods:** The European Union Labour Force Survey 2014 (11,345 women; 9,825 men) was used to study the immigrant working-age population (20-64 years) from seven West-European countries with a substantial number of refugees.

**Results:** A third had arrived as children (0-19 years). Each reason for migration was well represented within all country-groups, and the proportion of respondents reporting each reason was fairly similar across the country-groups. Regression analysis identified significant variation in education, employment and language-skills by reasons for migration within country-groups and vice versa, with (female) refugees and family-migrants arriving as adults faring worse than other migrants in language-skills and employment. There were few significant gender differences.

**Conclusion:** We recommend implementing reasons for migration and age upon arrival as core variables in quantitative migration studies.

**Keywords:** Reason for migration, Western Europe, refugees, immigrants, 1.5 generation, integration

## Introduction

Following large waves of immigrants to Western Europe in the 2000s, the economic and social integration of immigrants has become one of the most important areas of public debate and politics in Europe, and more research has been welcomed (OECD 2017; Font & Méndez 2013). However, research on immigrants within and between European countries faces a number of challenges, including the comparability and harmonisation of large-scale quantitative data sources, and how we classify and operationalize migrant groups, including refugees (Morales & Ros 2013; Sumption 2018; UNSC 2018).

There are many different typologies in the research literature on migrants. For example, country of birth can be used to classify immigrants as Western or non-Western, time spent in the host country identifies recently and less recently arrived immigrants, and entry and country residence can be regular or irregular. These typologies are related to cultural distance, the acculturation process and a migrant's civil and social rights, such as a residence or work permit and access to health care and other services (Samers & Collyer 2017). Consequently, how we theorise and operationalise migrants into typologies have consequences for how we describe and explain phenomena related to migrants, such as their economic and social integration, and for policy making relying on sound statistics and research.

A review of previous research on immigrant integration in Northern Europe has suggested that *age upon arrival* and *reason for migration* are two important theoretical dimensions that should be included in research aiming to explain the economic and social integration of immigrants in Western Europe (Pekkala & Kerr 2011). Key outcomes on integration are employment and the immigrants' human capital in the labour market, such as their education and language acquisition (Samers & Collyer 2017).

A large proportion of immigrants arrive as children or adolescents. To separate these children from first- and second-generation immigrants the term *generation one and a half* has been suggested (Rumbaut 2012). These children do not have the human capital or work experience that adult migrants carry with them, but arriving at a young age gives them the opportunity to socialise, learn the language and earn educational qualifications in the host-country before entering the labour market.

Immigrants migrate for a variety of reasons, both historical-structural, such as the enlargement of the European Union with an expanding common market with free movement (Samers & Collyer 2017), and motivational such as international protection (henceforth refugees), searching for work, studies, family reasons (for example children migrating with their parents, marriage and family reunions) and other reasons, such as health tourism. While refugees tend to stay on a permanent basis, labour-migrants, students and for example health tourists are likely to live trans-nationally mobile lives or return home once their objectives behind the migration have been achieved (Samers & Collyer 2017).

Male and female immigrants can be significantly different in their reasons for migration, the migration journey itself, and the integration process and its outcomes, such as employment (Pekkala & Kerr 2011), but depending on the context, also surprisingly similar. Consequently, to avoid a gender bias, we need to include a gender perspective both when theorising and reporting migration related phenomena (O' Reilly 2012; Samers & Collyer 2017).

## Surveying immigrants

The type of immigration and the kind of information sought should guide our choice of data collection. For example, a qualitative approach would be required to research irregular immigrants or circular trans-national mobility (Faist *et al.* 2013). However, for the majority of long-staying immigrants who reside on a legal basis with a registered household address, we can use more conventional methods to collect census data, administrative register data, survey data (Font & Méndez 2013), and in the future, big data (UNSC 2018).

Most national registers cover the entire population with the advantage that specific sub-groups of immigrants can be studied in detail and over time, but variables such as reason for migration and education obtained abroad can be missing. Furthermore, country-specific classifications often prevent comparisons (UNSC 2018). To cover such gaps, we need harmonised cross-country surveys. Eurostat's general population survey, the *European Labour Force Survey* (LFS), is the primary source for studying employment within and between European countries (Font & Méndez 2013).

Unfortunately, variables measuring reason for migration and age upon arrival are often excluded from general population surveys, such as the LFS. Consequently, immigrants are often classified into categories based on their country of birth, nationality and year of arrival (Sumption 2018), and statistical results may be interpreted on the assumption that migrants arriving in certain years from certain country-groups are strongly associated with specific reasons for migration, such as refugees (Gibson-Helm *et al.* 2014). This raises the issue how these variables are associated and what can be gained by utilising all three variables in combination.

### *Aims and hypotheses*

Our study has a methodological focus. It aims to explore the association between self-reported reasons for migration, age upon arrival and country-groups based on country of birth among the working-age immigrant population (age 20-64), and to study these three measures in combination in relation to education, employment and language skills. Substantial variations in integration outcomes for different reasons of migration within country groups, and for specific reasons such as refugees across country groups, would suggest that these measures should be implemented in future study designs.

We specify seven interrelated hypotheses that seek to elucidate the variation in reasons for migration across Eurostat's country-groups, and the need to identify refugees in particular when studying migrant integration: There is substantial variation in age upon arrival across reasons for migration (H1) and country-groups (H2), and specific reasons for migration should be highly associated with specific country-groups (H3). There is large variation in education, employment and language skills between reasons for migration within (H4) and between (H5) country-groups, with refugees standing out from the other reasons for migration (H6), and among the refugees and family-migrants the integration outcomes are better for those who arrived as children (H7). In addition, we report any notable gender differences.

## **Methods**

### *Data*

In 2014 the LFS included a module (AHM) that asked immigrants about their reasons for migration. Because the analysis focuses on integration outcomes, including employment, the target population is the working-age population (20 to 64 years old). We pooled together 9,825 male and 11,345 female immigrants from seven West-European

countries with a substantial proportion of refugees within each national sample (N; % refugees): Austria (3,974; 9.9%), Belgium (2,693; 9.5%), Finland (651; 8.5%), France (2,570; 4.5%), Norway (1,498; 13.1%), Sweden (2,110; 19.7%) and the United Kingdom (7,674; 4.7%). For more information about the survey see Eurostat (2017). The responsibility for all conclusions drawn from the data lies entirely with the authors. No table cells or predicted estimates were based on 3 or fewer observations.

### *Variables*

*Age upon arrival* was coded as 0-14, 15-19, 20-24, 25-29, 30-34, 35-39 and 40-64. *Main reason for migration* had six categories (figure labels): job, offered before migration (JO); job searcher (JS); family or marriage (FM); student (S); international protection or asylum (RA); other (O). *Age upon arrival* was used to extend this classification by identifying family (FC) and refugee (RC) children who arrived when 0-19 years old.

To harmonise Eurostat's anonymised *country of birth-classification* we recoded it as: EU-15 plus Australia, Canada, EFTA countries, New Zealand, Switzerland and the USA; EU28 and other European countries; North-Africa and the Middle East; Other Africa; Other Asia; Latin-America and Caribbean.

We dichotomized the three dependent variables used in the regression analysis. *Language skill* was recoded as mother tongue/advanced (0) versus intermediate/beginner (1), *employment* as working at least one hour per week (1) or not (0), and *educational level* as primary education (1) versus secondary/tertiary education (0).

Age (in 5-years intervals), parents' highest achieved educational level (primary, secondary or tertiary), host country and years lived in host country (from 1 to 11 or more) are dummy coded control variables.

### *Analysis*

We used binary logistic regression with the AHM-weights to estimate and visualise predicted scores of education (N=19.807), employment (N=19.807) and language skills (N=19.742) by country-groups and reasons for migration, using STATA 14.2. The tabulated results and descriptive statistics for all regressions are reported in the online appendix.

## **Results**

### *Age upon arrival by reasons for migration and country-groups*

Nearly three quarters of the immigrants arrived before the age of 30 (Table A2.1), and across the country-groups only 5-9% migrated when they were 40 to 64 years old (Table A1.2). Among family-migrants 45% were 14 years or younger, and 10% were 15 to 19 years old. Similar statistics for refugees were 1% and 13.5% (Table A1.1). This result suggests the 1.5 generation represents a large proportion of all adults of working-age in Western Europe.

Most labour-migrants and students arrived in their 20s, with more female than male students arriving as adolescents and more male students arriving in their late 20s. The distribution among refugees is more evenly spread out and gender equal than among the other groups, with a peak around 25 to 29 years of age. Only 11 percent of refugees arrived 40 years or older. Among family-migrants, more women than men arrived as adults. For all other age upon arrival intervals across reasons for migration there were only a few percentage points separating men and women.

Across the country-groups there were two notable differences. First, the proportion arriving at age 14 years or younger was higher in the EU15 and other Western countries (31%), North Africa and the Middle East (33%), and Latin America and the Caribbean (32%), than in Other Africa (22%), Other Asia (20%) and the EU28 and other Europe (14%), with a similarly reversed pattern for those aged 20 to 24. Second, the proportion arriving aged 15 to 19 (ranging from 10 to 14%), and the intervals for those who arrived 25 years or older, is strikingly similar across the country-groups. These data therefore suggest there are *more similarities than differences* in age upon arrival across country-groups and reasons for migration, with the main differences being the proportion arriving as children from various country-groups and adult women compared to adult men who arrived for family-reasons or work. These results do not support hypotheses one and two, which stated there should be substantial variation in age upon arrival across reasons for migration and country-groups.

#### *The association between reasons for migration and country-groups*

Table A1.3 presents the variation in reasons for migration across country-groups. For the four non-Western country-groups the overall pattern is fairly similar, with 5-13% refugees (including 0.5-2% of the total being children), 12-18% labour-migrants, 10-15% students, 21-30% adult family-migrants, 25.5-39% children family-migrants, and 4-7% who migrated for other reasons. Among the two Western country-groups a much larger proportion were labour or family-migrants, and 7% of the EU28 and other Europe immigrants were refugees. This figure is not different from the proportion who arrived from the non-Western countries. Moreover, the proportion of immigrants arriving from non-Western countries as refugees or asylum seekers is smaller than the proportion that arrived as either labour-migrants or students, and compared with the two Western country-groups, the proportion of student migrants is larger within all four non-Western country-groups.

However, these totals mask within-country-groups gender differences that were fairly similar across the four non-Western country-groups, such as more male than female students in three groups, and more male than female refugees and labour-migrants and more adult female family-migrants across all groups. For the other reasons for migration there were small gender differences. To conclude, no particular reason for migration is strongly associated with any particular country group, each reason for migration is well represented within most country groups, and there are similarities in the distribution of reasons for migration across the non-Western country-groups. Consequently, these results do not support hypothesis three.

[Figures 1-3 about here]

#### *Regression analysis*

Primary education, employment and having beginner/intermediate language-skills were regressed on reasons for migration and country-groups, while controlling for age, years of living in host country, country, gender and parental educational level (Table A3.1). Each model was extended with three two-way interaction-terms and the three-way interaction-term gender by migration-reasons by country-groups. All three-way interactions were statistically significant ( $p < 0.001$ ), and our visual presentation of the results are based on the models including all interactions. Predicted probabilities with 95% confidence intervals for each outcome by gender and reasons for migration across country-groups are presented in Figures 1 to 3.

In all three plots we see large variation in predicted probabilities by reasons for migration within country-groups. Across country-groups we also notice that the predicted score for a particular reason for migration, such as refugees, varies considerably. Without going into detail, these results support hypotheses four to seven. We would nevertheless like to point out that the variations, including some substantial gender differences, are larger with respect to employment and language skills, and within several country-groups the refugees and family-migrants who arrived as children had better predicted outcomes than migrants arriving as adults.

## Discussion

We did not find support for hypotheses one to three, which means that the age upon arrival variation across reasons for migration and country-groups is not large, and that no country group is highly associated with any particular reason for immigration, including where refugees are concerned. Consequently, country-group classifications should not be used as a crude proxy for reasons for migration to West-European countries (Gibson-Helm *et al.* 2014).

We found support for hypotheses four to seven. There was large variation in integration outcomes by reasons for migration within and between the country-groups. As expected, within most country-groups the refugees fared worse on all three integration outcomes, but in line with the 1.5 generation theory (Rumbaut 2012), refugees arriving as children often scored better.

Other than some systematic differences that were expected, such as more male labour-migrants and female family-migrants, or higher male employment rates, the results were generally more similar than different between men and women, which supports the feminist critique that male and female immigrants can be more similar than is often assumed (O' Reilly 2012; Samers & Collyer 2017).

Our results strongly support the conclusion that age upon arrival and reasons for migration are important theoretical dimensions that should be included in large-scale survey or register-based population studies aiming to explain the economic and social integration of immigrants (Pekkala & Kerr 2011). Future research should preferably include both measurements as complementary theoretical dimensions, reflecting the fact that migrants with similar reasons nevertheless can differ due to structural-historical conditions associated with their country of origin, their migration journey and where they settle. Furthermore, these combinations could be mediated by gender, meaning that male and female immigrants with similar origin and reasons for migration may nevertheless have different integration outcomes.

### *Methodological limitations*

The sampling frame for the LFS is the general population in each country (Eurostat 2017). Consequently, the proportion of immigrants and specific categories of immigrants, such as refugees, can be so small that reliable estimates cannot be made for certain countries. Moreover, we question how representative the data are for immigrants. Immigrants often settle in urban areas, and due to geographical segregation, the concentration of immigrants can be high in specific geographical locations. A sampling frame that does not acknowledge this difference between natives and immigrants could lead to systematic participation bias (Font & Méndez 2013; Samers & Collyer 2017; UNSC 2018). Interview language and whether the questionnaire exists in other languages, such as English, could be another barrier affecting

participation and create systematic differences between countries and immigrant groups within countries. Consequently, minorities are less likely than natives to participate in national representative surveys (Feskens *et al.* 2006; Font & Méndez 2013).

The study was cross-sectional, and thus without control for migrants who may have left the population. This could bias the results (Faist *et al.* 2013). Consequently, the results presented here are more likely to represent permanent, long-staying, educated and well-integrated immigrants living in residential areas mirroring that of the general population. Moreover, to have sufficiently large samples we pooled data from countries with different immigration histories and demographics, meaning country specific results could be different from ours.

To overcome these limitations, we need to rethink how we collect large-scale surveys of immigrants in Europe. As a first step we need national-representative samples of immigrants with sufficiently large sample sizes to identify and study refugees and other types of immigrants within and between countries, and a panel design would further strengthen the scientific quality (Font & Méndez 2013). Despite its costs, without such investments comparable European studies with a high scientific impact cannot be carried out, deeming it impossible to produce the knowledge stakeholders desire and need for policy making.

## **Conclusion**

In the regression analyses there was significant variation in integration outcomes between reasons for migration both within and between country-groups. This result suggests these two theoretical dimensions are independent of each other, and as shown in the cross-tabulated association between the two measures, country-groups are a poor proxy measurement for reasons for migration.

A third of the working-age immigrants had arrived as children, and refugees and family-migrants who arrived as children were better integrated than those who arrived as adults, suggesting that the 1.5-generation acculturation theory is important. We therefore strongly recommend that reasons for migration and age upon arrival are implemented as core variables in quantitative immigrant studies, and researchers should whenever possible combine both with country-groups in their study designs.

There are currently few large-scale data available to study immigrants within and between European countries. Our usage of the best available data source suggests there is still a long way to go before high-quality harmonised data are available to researchers. For both research and policy purposes these issues need to be solved and implemented in large-scale quantitative studies.

## **Acknowledgement and funding**

TBC

## **Conflict of interest**

The authors have no conflict of interest to report.

## **Literature**

**EUROSTAT** (2017) *EU Labour Force Survey Database User Guide, Version September 2017* Luxembourg: European Commission, Eurostat

**FAIST, T., FAUSER, M. and REISENAUER, E.** (2013) *Transnational migration*. Cambridge: Polity Press

**FESKENS, R., HOX, J., LENSVELT-MULDERS, G. and SCHMEETS, H.** (2006) 'Collecting data among ethnic minorities in an international perspective'. *Field Methods*, 18(3): 284-304

**FONT, J. and MÉNDEZ, M.** (2013) 'Introduction: The methodological challenges of surveying populations of immigrant origin'. In Font, J and Méndez, M (eds.) *Surveying ethnic minorities and immigrant populations: Methodological challenges and research strategies*. Amsterdam: Amsterdam University Press

**GIBSON-HELM, M., BOYLE, J., BLOCK, A. and TEEDE, H.** (2014) 'Use of country of birth as an indicator of refugee background in health datasets'. *BMC Medical Research Methodology*, 14:27

**MORALES, L. and ROS, V.** (2013) 'Comparing the response rates of autochthonous and migrant populations in nominal sampling surveys: The LOCALMULTIDEM study in Madrid'. In Font, J. and Méndez, M. (eds.) *Surveying ethnic minorities and immigrant populations: Methodological challenges and research strategies*. Amsterdam: Amsterdam University Press

**OECD** (2017) '*International Migration Outlook*', Paris: OECD Publishing.  
[http://doi.org/10.1787/migr\\_outlook-2017-en](http://doi.org/10.1787/migr_outlook-2017-en) (accessed 2 May2019)

**PEKKALA KERR, S. and KERR, W.** (2011) 'Economic impacts of immigration: A survey'. *NBER Working Paper Series*, Working Paper No. 16736

**RICHMOND, A.H.** (1988) 'Sociological theories of international migration: The case of refugees'. *Current Sociology*, 36(2): 7-25

**RUMBAUT, R. G.** (2012) 'Generation 1.5, Educational experiences of.' In Banks, J.A. *Encyclopedia of diversity in education*. Thousand Oaks, CA: Sage Publications

**SAMERS, M. and COLLYER, M.** (2017) *Migration*. London: Routledge

**SUMPTION, M.** (2018) 'Reason for migration statistics and policy research'. In *Power from statistics: data, information and knowledge*. Eurostat

**UNITED NATIONS STATISTICAL COMMITTEE** (2018) *International recommendations on refugee statistics*. The UN Refugee Agency